Rasch-Modelle und Verallgemeinerung

Seminar: Ausgewählte Aspekte der Wirtschafts- und Sozialstatistik

Nataliia Semenenko

Ludwig-Maximillian-Universität München
Institut für Statistik

München, 22. Mai 2015

Inhaltsverzeichnis I

Inhaltsverzeichnis II

Einleitung

Das Rasch-Modell:

- wurde von dänischen Statistiker Georg Rasch entwickelt (1960)
- das Grundmodell aller Modelle der *Item-Response-Theorie* (**IRT**)

Ziel:

- beobachtete Reaktion durch die latente Eigenschaft zu erklären
- Schätzung der nicht beobachteten Eigenschaften

Berühmtester Einsatz: PISA-Studie der OECD

n - Anzahl Personen

m - Anzahl Aufgaben

$u \in \{0{,}1\}$ - beobachtete Variable Aufgabe gelöst ($1 =$ ja, $0 =$ nein)

Beispiel (für n = 7, m = 7)

	Aufgabe						
Person	*1*	*2*	*3*	*4*	*5*	*6*	*7*
1	*0*	*1*	*0*	*1*	*0*	*0*	*0*
2	*0*	*1*	*1*	*1*	*0*	*1*	*0*
3	*1*	*1*	*0*	*0*	*0*	*1*	*1*
4	*0*	*0*	*0*	*0*	*1*	*1*	*0*
5	*0*	*0*	*0*	*0*	*0*	*1*	*0*
6	*1*	*1*	*1*	*1*	*1*	*1*	*0*
7	*1*	*0*	*0*	*0*	*1*	*0*	*0*

Allgemeine Form

$i = 1, ...n$ - Laufindex für Personen, die an einem Test teilnehmen
$j = 1, ...m$ - Laufindex für Aufgaben im Test
$u_{i,j} \in \{0{,}1\}$ - beobachtete Variable Aufgabe gelöst (1 = ja, 0 = nein)
Eintrag für die i-te Person und j-te Aufgabe

					Aufgabe		
Person	1	2	3	...	**j**	...	m
1	$u_{1,1}$	$u_{1,2}$	$u_{1,3}$	...	$u_{1,j}$	...	$u_{1,m}$
2	$u_{2,1}$	$u_{2,2}$			.		.
3	$u_{3,1}$		.		.		.
.	.			.	.		.
i	$u_{i,1}$	.	.	.	$u_{i,j}$		.
.						.	.
n	$u_{n,1}$	.	.	...	.	...	$u_{n,m}$

Grundform des Rasch-Modells

Modellgleichung

θ - Personen-Parameter, z.B. Fähigkeit einer Person

β - Aufgaben-Parameter, z.B. Schwierigkeit einer Aufgabe

$U_{i,j} \in \{0,1\}$ - unbekannte Variable Aufgabe gelöst ($1 =$ ja, $0 =$ nein)
Eintrag für die i-te Person und j-te Aufgabe

$\mathrm{P}(U_{i,j} = u_{i,j})$ - Lösungswahrscheinlichkeit.

Forderungen

1. Lösungswahrscheinlichkeit abhängig von der Fähigkeit der Person θ_i
2. Lösungswahrscheinlichkeit abhängig von der Schwierigkeit der Aufgabe β_j
3. positiver Zusammenhang zwischen der Personenfähigkeit und Lösungswahrscheinlichkeit
4. $P(U_{i,j} = u_{i,j}) \in [0,1]$

Forderungen

1. Lösungswahrscheinlichkeit abhängig von der Fähigkeit der Person θ_i
2. Lösungswahrscheinlichkeit abhängig von der Schwierigkeit der Aufgabe β_j
3. positiver Zusammenhang zwischen der Personenfähigkeit und Lösungswahrscheinlichkeit
4. $P(U_{i,j} = u_{i,j}) \in [0,1]$

Rasch-Modell

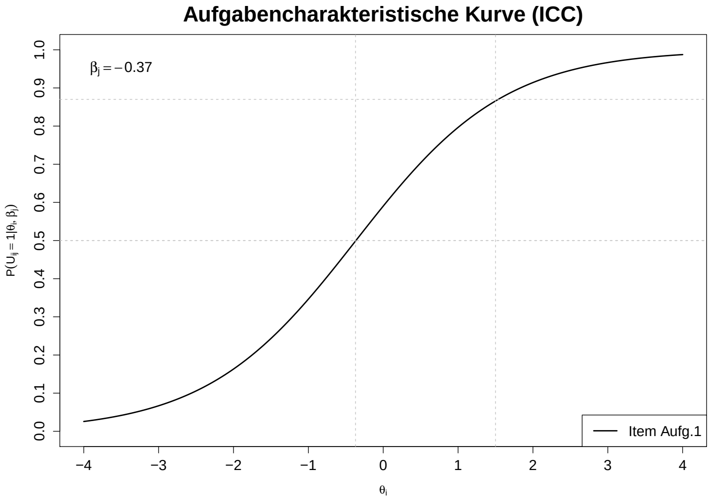
$$\boxed{P(U_{i,j} = 1 | \theta_i, \beta_j) = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}}}$$

Logit Modell

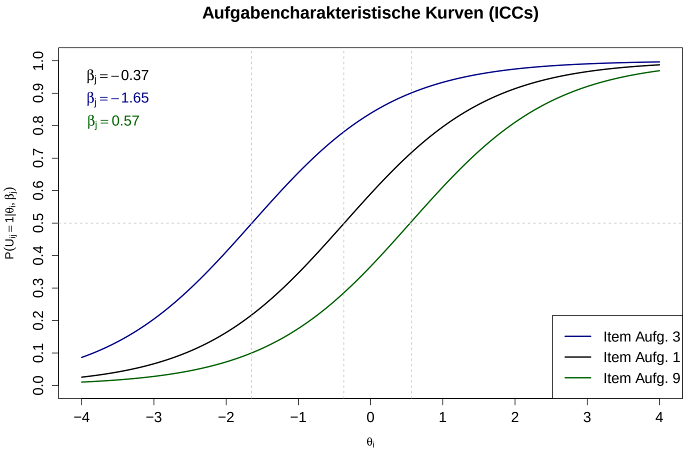
mit ζ_i - der lineare Prädiktor

$$\boxed{P(Y_i = 1) = \pi_i = \frac{e^{\zeta_i}}{1 + e^{\zeta_i}}}$$

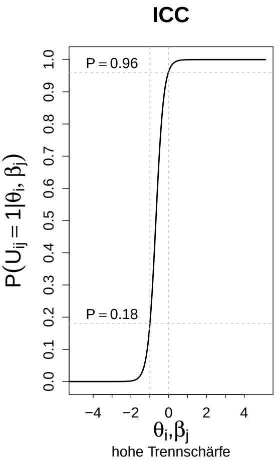
ICC (*Item Charasteristic Curve*)

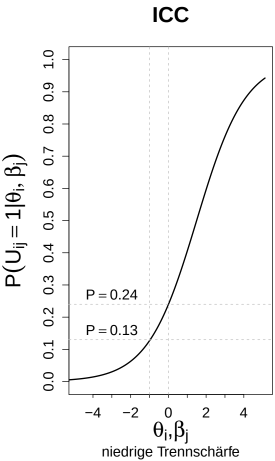


ICCs (*Item Charasteristic Curves*)



ICC (*Item Charasteristic Curve*)





Zentrale Annahmen und Eigenschaften

- Suffiziente Statistiken
- Lokale Stochastische Unabhängigkeit
- Spezifische Objektivität
- Eindimensionalität

Ziel: Alle unbekannte Parameter ($\theta_i$undβ_j) zu schätzen

Methode: Maximum-Likelihood-Schätzung

Möglichkeiten:

- $\theta_i und \beta_j$ gleichzeitig geschätzt (Gemeinsame ML-Schätzung)
- $\theta_i und \beta_j$ nacheinander geschätzt (Bedingte ML-Schätzung oder Marginale ML-Schätzung)

Die Aufgaben-Parameter sind als Linearkombination der Parameter für die Teilkompetenzen dargestellt:

$$\beta_j = \sum_l \omega_{j,l} \cdot \eta_l$$

mit

- $l = 1, ..., L$ - Laufindex für nötige Teilkompetenzen für eine Aufgabe
- $\omega_{j,l} \in \{0,1\}$ - Gewicht, Teilkompetenz in der Aufgabe enthalten (1=ja, 0=nein)
- η_l - Schwierigkeit der Teilkompetenz

Verallgemeinerungen des Rasch-Modells

Das linear-logistische Testmodell (LLTM)

Gewichtungstabelle

	Teilkompetenzen						
Aufgabe	1	2	3	...	**l**	...	L
1	$\omega_{1,1}$	$\omega_{1,2}$	$\omega_{1,3}$	...	$\omega_{1,j}$	...	$\omega_{1,L}$
2	$\omega_{2,1}$	$\omega_{2,2}$			.		.
3	$\omega_{3,1}$		.		.		.
.	.			.	.		.
j	$\omega_{j,1}$	.	.	.	$\omega_{j,l}$		.
.						.	.
m	$\omega_{m,1}$	.	.	...	.	...	$\omega_{m,L}$

Beispiel (Gewichtungstabelle für 3 Aufgaben, 3 Teilkompetenzen)

Aufgabe	**Teilkompetenzen** *Kommutativ-Gesetz*	*Assoziativ-Gesetz*	*Distributiv-Gesetz*
1	*0*	*1*	*1*
2	*1*	*1*	*0*
3	*0*	*0*	*1*

Das zwei-parametrige Birnbaum-Modell (2PLM)

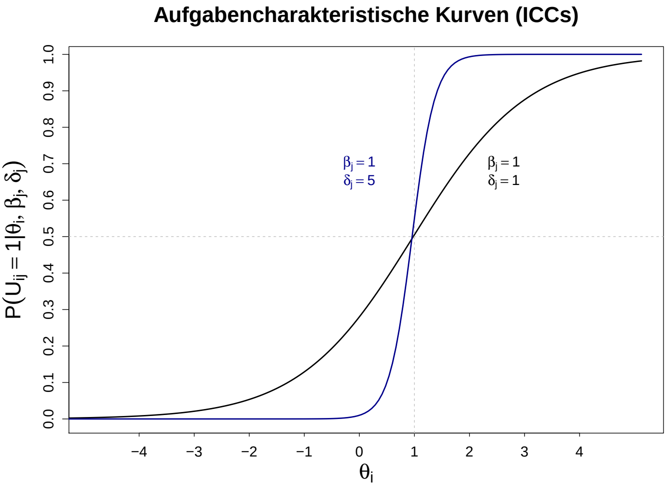
δ_j - Diskriminationsparameter,Steigung j-ter Aufgabe

bisher: $\delta_j = 1$
jetzt: $\delta_j \neq 1$

$$\boxed{P(U_{i,j} = 1 | \theta_i, \beta_j, \delta_j) = \frac{e^{\delta_j(\theta_i - \beta_j)}}{1 + e^{\delta_j(\theta_i - \beta_j)}}}$$

$\Rightarrow$ keine spezifische Objektivität

Das zwei-parametrige Birnbaum-Modell (2PLM)



Das Birnbaum-Modell mit zusätzlichem Rateparameter (3PLM)

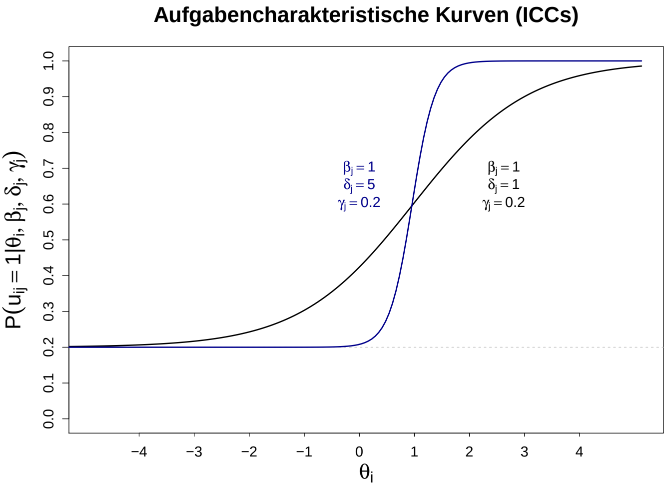
δ_j - Steigung j-ter Aufgabe

γ_j - Regulierungs-Parameter für Lösungswahrscheinlichkeiten

bisher: $\gamma_j = 0$
jetzt: $\gamma_j \neq 0$

$$\boxed{P(U_{i,j} = 1 | \theta_i, \beta_j, \delta_j, \gamma_j) = \gamma_j + (1 - \gamma_j) \cdot \frac{e^{\delta_j(\theta_i - \beta_j)}}{1 + e^{\delta_j(\theta_i - \beta_j)}}}$$

Das Birnbaum-Modell mit zusätzlichem Rateparameter (3PLM)



$\beta_{j,c}$ - Parameter für c-te Antwortkategorie der j-ten Aufgabe

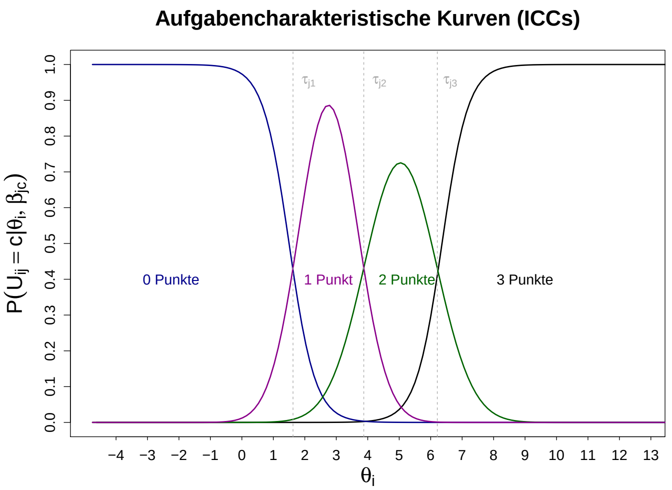
$\tau_{j,l}$ - Schwellenwerte

bisher: $u_{i,j} \in \{0{,}1\}$ - Variable Aufgabe gelöst bzw. Aussage zugestimmt (1 = ja, 0 = nein)
jetzt: $u_{i,j} \in \{0,...,m_j\}$ - Variable Aufgabe gelöst bzw. Aussage zugestimmt (0 = nein, 1 = 1.Teilschritt, 2 = 1. und 2. Teilschritte,...)

$$\boxed{P(U_{i,j} = c|\theta_i, \beta_j) = \frac{e^{c\cdot\theta_i - \beta_{j,c}}}{\sum_{l=0}^{m_j} e^{l\cdot\theta_i - \beta_{j,l}}}}$$

mit $\beta_{j,c} = \sum_{l=1}^{c} \tau_{j,l}$ und $\beta_{j,0} = 0$

Beispiel (4 Kategorien: 0, 1, 2 oder 3 Punkte)



Verallgemeinerungen des Rasch-Modells

weitere Verallgemeinerungen

- Das Raiting-Scale-Modell
- Das dichtome Rasch-Modell mit Rateparametern
- Das Mischverteilungs-Rasch-Modell (Mixed-Rasch-Modell)
- Die Mehrdimensionale Rasch-Modelle
- u.a.

Modellgeltungstests

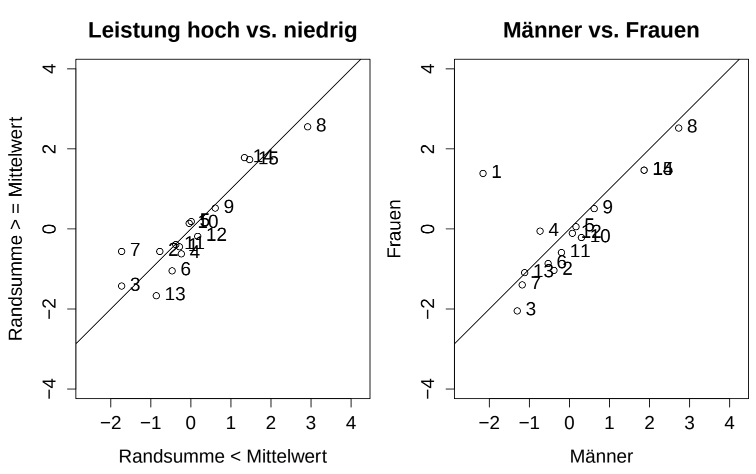
- Der Graphische Modelltest
- Likelihood-Quotienten-Test
- Wald-Tests

Idee:

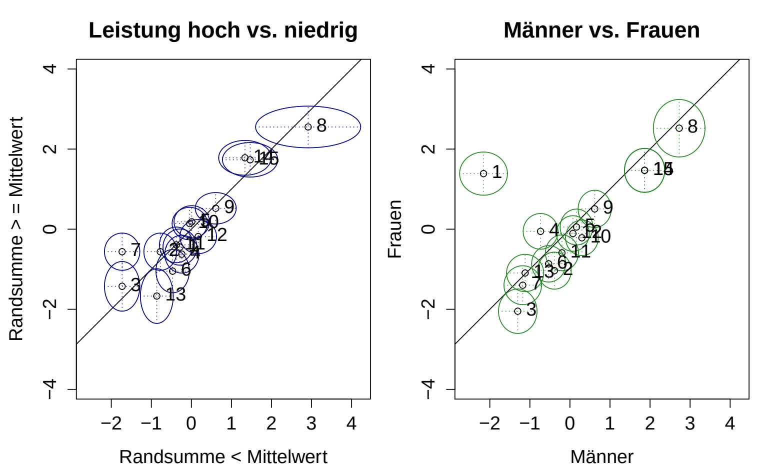
- **keine systematische Unterschiede** in der geschätzten Aufgaben-Parameter zw. den Gruppen ⇒ **Rasch-Modell gilt**
- **DIF** (*Differential Item Functioning*)⇒ **Rasch-Modell gilt nicht**

Idee: Vergleiche geschätzte Aufgaben-Parameter in 2 Gruppen.
Rasch-Modell gilt $\Rightarrow$ Werte in beiden Gruppen stimmen überein (bis auf lineare Transformationen)

Beispiel (Simulierte Daten: 200 Personen, 15 Aufgaben)



Beispiel (Simulierte Daten: 200 Personen, 15 Aufgaben)



- $k = 1...K$ - Laufindex für den Gruppen
- β_k - Aufgaben-Parameter aus einzelne Gruppe
- β - Aufgaben-Parameter aus Gesamtstichprobe

$$H_0 : LQ = 1 \text{ vs. } H_1 : LQ \neq 1$$

mit:

$$LQ = \frac{L_u(r, \hat{\beta})}{\prod_{k=1}^{K} L_{u_k}(r_k, \hat{\beta}_k)}$$

Teststatistik:

$$T = -2 \cdot ln(LQ) \sim \chi^2(K-1) \cdot (m-1)$$

H_0 abzulehnen, falls: $T > z_{1-\alpha}(\chi^2)$

Leistung

Andersen LR-test:	
LR-value:	24.807
Chi-square df:	14
p-value:	**0.037**

Geschlecht

Andersen LR-test:	
LR-value:	111.721
Chi-square df:	14
p-value:	**0**

Schultyp

Andersen LR-test:	
LR-value:	49.155
Chi-square df:	42
p-value:	**0.208**

- **Leistung:**
  - hoch
  - niedrig
- **Geschlecht:**
  - Männer
  - Frauen
- **Schultyp:**
  - Gymnasium
  - Realschule
  - Hauptschule
  - Integrierte Gesamtschule

$$H_0 : \hat{\beta}_{j,k} = \hat{\beta}_{j,l} \text{ vs.} H_1 : \hat{\beta}_{j,k} \neq \hat{\beta}_{j,l} \text{ für } k \neq l$$

$$W_j = \frac{(\hat{\beta}_{j,1} - \hat{\beta}_{j,2})^2}{\hat{\sigma}^2_{j,1} + \hat{\sigma}^2_{j,2}}$$

Teststatistiken:

$$T = sign(\hat{\beta}_{j,1} - \hat{\beta}_{j,2}) \cdot \sqrt{W_j} \sim N(0,1)$$

$$T = (\hat{\beta}_1 - \hat{\beta}_2)'(\hat{\Sigma}_1 - \hat{\Sigma}_2)^{-1}(\hat{\beta}_1 - \hat{\beta}_2) \sim \chi^2(m-1)$$

H_0 abzulehnen, falls: $T > z_{1-\alpha}(N \text{ bzw. } \chi^2)$

Modellgeltungstests

Wald-Test

Leistung hoch vs. niedrig

Wald test on item level (z-values):		
	z-statistic	p-value
beta Aufg. 1	-0.512	0.609
beta Aufg. 2	0.677	0.499
beta Aufg. 3	0.780	0.435
beta Aufg. 4	-1.186	0.236
beta Aufg. 5	0.573	0.567
beta Aufg. 6	-1.667	0.095
beta Aufg. 7	3.570	**0.000**
beta Aufg. 8	-0.504	0.614
beta Aufg. 9	-0.262	0.794
beta Aufg. 10	0.599	0.549
beta Aufg. 11	-0.044	0.965
beta Aufg. 12	-1.114	0.265
beta Aufg. 13	-1.986	**0.047**
beta Aufg. 14	1.107	0.268
beta Aufg. 15	0.640	0.522

Männer vs. Frauen

Wald test on item level (z-values):		
	z-statistic	p-value
beta Aufg. 1	8.643	**0.000**
beta Aufg. 2	-2.039	**0.041**
beta Aufg. 3	-1.991	**0.046**
beta Aufg. 4	2.139	**0.032**
beta Aufg. 5	-0.354	0.723
beta Aufg. 6	-1.049	0.294
beta Aufg. 7	-0.638	0.523
beta Aufg. 8	-0.420	0.675
beta Aufg. 9	-0.351	0.725
beta Aufg. 10	-1.672	0.095
beta Aufg. 11	-1.260	0.208
beta Aufg. 12	-0.590	0.555
beta Aufg. 13	0.076	0.939
beta Aufg. 14	-1.041	0.298
beta Aufg. 15	-1.041	0.298

Zusammenfassung und Ausblick

- Rasch-Modell und Verallgemeinerungen spielen grundlegende Rolle in der **IRT**
- Testentwicklung $\Rightarrow$ Rasch-Modell
- Analysen von durchgeführten Tests $\Rightarrow$ Verallgemeinerungen

Weitere Fragestellungen:

- Weitere Verallgemeinerungen detailliert
- χ^2- Anpassungstest
- Analyse anhand Simulationen von verallgemeinerten Modellen

Vielen Dank für Ihre Aufmerksamkeit!

Quellenverzeichnis

Strobl, C. (2010).
Das Rasch-Modell: Eine verständliche Einführung für Studium und Praxis.
Band 2, Rainer Hamp Verlag, München und Mering

Kubinger, K.D., Rasch, D., Yanagida, T. (2011).
Statistik in der Psychologie. Von Einführungskurs bis zur Dissertation. Lehrbuch.
Hogrefe Verlag, S. 555-568.

Moosbrugger, H., Kelava, A. (2012).
Testtheorie und Fragebogenkonstruktion.Mit 66 Abbildungen und 41 Tabellen.
2., aktualisierte und überarbeitete Auflage, Springer-Verlag Berlin Heidelberg, S. 227-274.

Quellenverzeichnis

Henning, H.J. (1974).
Skalenanalyse und Rasch-Modell.
Bonn

Penzel, M., Sälzer, C., Klieme, E., Köller, O.(Hrsg) (2013).
PISA 2012. Fortschritte und Herausforderungen in Deutschland.
Waxmann Verlag, S. 334-337.

Toutenburg, H.,
mit Beiträgen von Heumann, C., Nittner, T., Scheid, S. (2003).
Lineare Modelle. Theorie und Anwendungen.
2., neu bearbeitete und erweiterte Auflage, Physica Verlag, S. 402-403.

http://www.jstatsoft.org/

http://www.sozialwissenschaftliche-forschungsmethoden.de/

http://www.r-project.org/

Suffiziente Statistiken

- suffiziente Statistik $\neq$ erwartungstreue Statistik

Beispiel (für n=5:)

- $\bar{x} = \frac{1}{5}\sum_{i=1}^{5} x_i = \frac{1}{5}(x_1 + x_2 + x_3 + x_4 + x_5)$
  - *erwartungstreu* ✓
  - *suffizient* ✓
- $x^* = \frac{1}{3}(x_1 + x_3 + x_5)$
  - *erwartungstreu* ✓
  - **nicht suffizient**

Suffiziente Statistiken

Beispiel (für n = 7, m = 7)

	Aufgabe							
Person	1	2	3	4	5	6	7	r_i
1	0	1	0	1	0	0	0	2
2	0	1	1	1	0	1	0	4
3	1	1	0	0	0	1	1	4
4	0	0	0	0	1	1	0	2
5	0	0	0	0	0	1	0	1
6	1	1	1	1	1	1	0	6
7	1	0	0	0	1	0	0	2
s_j	3	4	2	3	3	5	1	

Lokale Stochastische Unabhängigkeit

Lösungswahrscheinlichkeit für **alle** Aufgaben einer Person:

$$P(U_{i,1} = u_{i,1}, ...U_{i,m} = u_{i,m} | \theta_i, \beta_1, ..., \beta_m) =$$

$$\prod_{j=1}^{m} P(U_{i,j} = u_{i,j} | \theta_i, \beta_j) =$$

$$\prod_{j=1}^{m} \frac{e^{u_{i,j}(\theta_i - \beta_j)}}{1 + e^{\theta_i - \beta_j}} =$$

$$\frac{e^{r_i \theta_i - \sum_{j=1}^{m} u_{i,j} \beta_j}}{\prod_{j=1}^{m} (1 + e^{\theta_i - \beta_j})}$$

Lokale Stochastische Unabhängigkeit

Lösungswahrscheinlichkeit für **alle** Aufgaben und **alle** Personen:

$$P(U_1 = u_1, ...U_n = u_n | \theta_1, ..., \theta_n, \beta) =$$

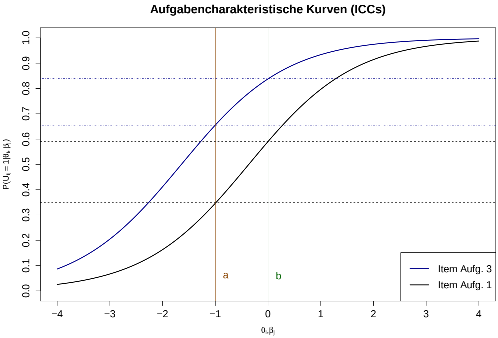
$$\prod_{i=1}^{n} \prod_{j=1}^{m} P(U_{i,j} = u_{i,j} | \theta_i, \beta_j) =$$

$$\prod_{i=1}^{n} \frac{e^{r_i \theta_i - \sum_{j=1}^{m} u_{i,j} \beta_j}}{\prod_{j=1}^{m} (1 + e^{\theta_i - \beta_j})} =$$

$$\frac{e^{\sum_{i=1}^{n} r_i \theta_i - \sum_{j=1}^{m} s_j \beta_j}}{\prod_{i=1}^{n} \prod_{j=1}^{m} (1 + e^{\theta_i - \beta_j})}$$

Spezifische Objektivität

- Aussagen über die Fähigkeiten von 2 Personen unabhängig von der gewählten Aufgabe
- Aussagen über die Schwierigkeit von 2 Aufgaben unabhängig von den gewählten Personen



Schritt 1: Schätzung der Aufgabenparameter β_j

Schritt 2: Schätzung der Personenparameter θ_i

Die Likelihood (für eine Person und **alle** Aufgaben):

$$L_{u_i}(\theta_i, \beta) = P(u_i|\theta_i, \beta) = \frac{e^{r_i\theta_i - \sum_{j=1}^{m} u_{i,j}\beta_j}}{\prod_{j=1}^{m}(1 + e^{\theta_i - \beta_j})}$$

$$L_{u_i}(\theta_i, \beta) = g(r_i|\theta_i, \beta) \cdot h(u_i|r_i, \theta_i, \beta)$$

$$\Leftrightarrow$$

$$h(u_i|r_i, \theta_i, \beta) = \frac{L_{u_i}(\theta_i, \beta)}{g(r_i|\theta_i, \beta)}$$

mit

$h(u_i|r_i, \theta_i, \beta)$ - bedingte auf r_i Likelihood für i-te Person

$g(r_i|\theta_i, \beta)$ - Wahrscheinlichkeit, eine bestimmte r_i zu beobachten

$$g(r_i|\theta_i,\beta) = \sum_{\sum_j u_{i.j}=r_i} P(u_i|\theta_i,\beta) =$$

$$\sum_{\sum_j u_{i.j}=r_i} \frac{e^{r_i\theta_i - \sum_{j=1}^m u_{i,j}\beta_j}}{\prod_{j=1}^m (1+e^{\theta_i-\beta_j})} =$$

$$\frac{e^{r_i\theta_i}}{\prod_{j=1}^m (1+e^{\theta_i-\beta_j})} \cdot \sum_{\sum_j u_{i.j}=r_i} e^{-\sum_{j=1}^m u_{i,j}\beta_j} =$$

$$\frac{e^{r_i\theta_i}}{\prod_{j=1}^m (1+e^{\theta_i-\beta_j})} \cdot \gamma_{r_i}(\beta)$$

Bedingte Likelihood für i-te Person:

$$h(u_i|r_i,\theta_i,\beta) = \frac{L_{u_i}(\theta_i,\beta)}{g(r_i|\theta_i,\beta)} =$$

$$\frac{e^{r_i\theta_i - \sum_{j=1}^{m} u_{i,j}\beta_j}}{\prod_{j=1}^{m}(1+e^{\theta_i-\beta_j})} : \frac{e^{r_i\theta_i}\gamma_{r_i}(\beta)}{\prod_{j=1}^{m}(1+e^{\theta_i-\beta_j})} =$$

$$\frac{e^{r_i\theta_i - \sum_{j=1}^{m} u_{i,j}\beta_j}}{\prod_{j=1}^{m}(1+e^{\theta_i-\beta_j})} \cdot \frac{\prod_{j=1}^{m}(1+e^{\theta_i-\beta_j})}{e^{r_i\theta_i}\gamma_{r_i}(\beta)} =$$

$$\frac{e^{-\sum_{j=1}^{m} u_{i,j}\beta_j}}{\gamma_{r_i}(\beta)} = h(u_i|r_i,\beta)$$

Bedingte Likelihood für gesamte Daten:

$$h(u|r,\beta) = \prod_{i=1}^{n} h(u_i|r_i,\beta) =$$

$$\prod_{i=1}^{n} \frac{e^{-\sum_{j=1}^{m} u_{i,j}\beta_j}}{\gamma_{r_i}(\beta)} =$$

$$\frac{e^{-\sum_{i=1}^{n}\sum_{j=1}^{m} u_{i,j}\beta_j}}{\prod_{i=1}^{n}\gamma_{r_i}(\beta)} =$$

$$\frac{e^{-\sum_{j=1}^{m} s_j\beta_j}}{\prod_{i=1}^{n}\gamma_{r_i}(\beta)}$$

Schritt 1: Schätzung der Aufgabenparameter β_j

Schritt 2: Schätzung der Personenparameter θ_i

Unterschied zu bedingten ML-Schätzung: Personenparameter werden rausintegriert.
Nötig: Annahme über marginale Verteilung (Randdichte $f(\theta)$ ∼N(0,1))
Problem: NV-Annahme für Randdichte $f(\theta)$ falsch → Schätzung verzerrt

$$L_u(\theta, \beta) \cdot f(\theta) = P(u|\theta, \beta) \cdot f(\theta) = P(u, \theta|\beta)$$

Marginale Likelihood für die Aufgabenparameter:

$$L_u(\beta) = \int P(u, \theta|\beta)\partial\theta$$